



УДК 004.522

**ОБ ОДНОМ ПОДХОДЕ К УМЕНЬШЕНИЮ ОБЪЕМА БИТОВЫХ
ПРЕДСТАВЛЕНИЙ РЕЧЕВЫХ ДАННЫХ БЕЗ ПАУЗ****ABOUT ONE APPROACH TO REDUCE THE VOLUME OF BIT VOTE
REPRESENTATIONS WITHOUT PAUSE****С.П. Белов¹, А.А. Медведева², А.В. Болдышев²
S.P. Belov¹, A.A. Medvedeva², A.V. Boldyshev²**¹⁾ Белгородский университет кооперации, экономики и права,
Россия, 308023, г. Белгород, ул. Садовая, д. 116а²⁾ Белгородский государственный национальный исследовательский университет,
Россия, 308015, г. Белгород, ул. Победы, д. 85¹⁾ Belgorod University of Cooperation, Economics and Law,
116a Sadovaya St., Belgorod, 308023, Russia²⁾ Belgorod State National Research University, 85 Pobeda St., Belgorod, 308015, RussiaE-mail: belov@bsu.edu.ru, medvedeva_aa@bsu.edu.ru, boldyshev@bsu.edu.ru**Аннотация**

Непрерывное увеличение в информационно-телекоммуникационных системах (ИТС) потоков речевых сообщений, как наиболее естественной для человека формы информационного обмена, привело к почти полной занятости частотно-временных ресурсов современных ИТС. В связи с этим, сегодня созданы методы уменьшения объемов битовых представлений речевых данных (сжатие) как за счет удаления пауз между отдельными словами, так и за счет сжатия речевых сигналов, порождаемых собственно звуками речи, что позволило существенно уменьшить затраты частотно-временных ресурсов ИТС при передаче этого вида трафика. Однако существующие сегодня подходы к сжатию собственно звуков речи, основанные на использовании психоакустической модели с применением грубого квантования по уровню так называемых субполосных преобразований отрезков (векторов) отсчетов речевых сигналов не являются оптимальными.

В работе предлагается метод сжатия речевых данных без пауз, созданный на основе применения математического аппарата собственных векторов субполосных матриц, позволяющего адекватно сформулировать вариационные условия и решить оптимизационные задачи обработки речевых данных.

Abstract

The continuous increase the flows of voice messages in the information and telecommunications systems (ITS), as the most natural form of information exchange for a person, has led to the almost full employment of the time-frequency resources of modern ITS. In this regard, today, methods have been created for reducing the amount of bit representations of speech data (compression), both by removing pauses between individual words, and by compressing speech signals generated by the actual speech sounds, which significantly reduced the cost of ITS time-frequency resources this type of traffic. However, existing approaches to compression only of the sounds of speech, based on the use of a psychoacoustic model with the use of coarse quantization by level of so-called subband transformations of segments of speech signal samples (vectors) are not optimal.

The paper proposes a method of compression of speech data without pauses, created on the basis of the application of the mathematical apparatus of the eigenvectors of subband matrices, which allows adequately to formulate variational conditions and solve optimization problems of processing speech data.

Ключевые слова: отрезок речевого сигнала, речевые данные, распределение энергии, субполосная матрица, собственные вектора субполосной матрицы, информационные частотные интервалы, коэффициент сжатия, кодовые книги квазиоптимальных квантователей.



Keywords: speech signal segment, speech data, energy distribution, subband matrix, eigenvectors of the subband matrix, information frequency intervals, compression ratio, codebooks of quasi-optimal quantizers.

Введение

Проблема уменьшения объемов битовых представлений речевых данных при их хранении и передаче рассматривается в работах многих авторов, особенно специалистов в области телекоммуникаций, что подтверждается результатами анализа научно-технической литературы [Сергиенко, Баринов, 2009.; Свириденко, 2009; Жиляков и др., 2007; Жиляков и др., 2005; Сэлмон, 2004.; Санников, 2003; Шелухин, 2000; Ashwin, Kumaresan, 2000; Seki, 1958; Калинин, 1991; Жиляков и др., 2010; Быков и др., 2003; Алдошина, 2002].

При этом отмечаются два основных аспекта: необходимость обнаружения с последующим их кодированием пауз [Шелухин, 2000; Калинин, 1991; Жиляков и др., 2010; Жиляков и др., 2007; Белов и Прохоренко, 2006; Быков и др., 2003], возникающих между отдельными словами и в режиме диалога занимающих до 60% длительности исходных звукозаписей, и сокращение объемов битовых представлений собственно звуков речи без пауз [Сергиенко, Баринов, 2009.; Свириденко, 2009; Алдошина, 2002; Ковалгин, Вологдин, 2004].

Существующие методы сжатия звуков речи без пауз с использованием грубого квантования по уровню основываются на психоакустической модели, что приводит к необходимости применения так называемых субполосных преобразований отрезков (векторов) отсчетов речевых сигналов, позволяющих получить другие векторы, подвекторы которых отражают частотные свойства исходного вектора в выбранных диапазонах оси частот [Жиляков и др., 2007; Гусинская, Зайцев, 2004; Сергиенко, 2003; Синильников, 1988]. Именно компоненты этих подвекторов подвергаются квантованию по уровню с различными шагами, чем достигается учет частотно-избирательных свойств человеческого слуха. В настоящее время для субполосного преобразования принято использовать процедуру прореживания выходных последовательностей КИХ-фильтров (фильтров с конечной импульсной характеристикой), настроенных на соответствующие участки оси частот [Сергиенко, 2003; Синильников, 1988]. Такая процедура субполосного преобразования не является оптимальной в смысле минимума погрешностей аппроксимации спектров исходных векторов в выбранных частотных диапазонах, что приводит к увеличению погрешностей восстановления данных по квантованным значениям и, как следствие, к ухудшению качества воспроизводимой речи [Синильников, 1988].

В связи с этим, авторами предлагается метод сжатия речевых данных без пауз, созданный на основе применения математического аппарата собственных векторов субполосных матриц, позволяющего адекватно сформулировать вариационные условия и решить оптимизационные задачи обработки речевых данных.

Математические основы метода

Осуществляется обработка отдельных отрезков (векторов отсчетов) сигнала речи:

$$\vec{x} = (x_1, \dots, x_N)^T \quad (1)$$

в соответствии с выбранным равномерным разбиением полосы нормированных частот $0 \leq \omega \leq \pi$ на R интервалов V_r вида:

$$V_{r2} - V_{r1} = \pi / R; \quad V_{r1+1} = V_{r2}; \quad r = 1; 2 \dots R-1 \quad (2)$$

одинаковой ширины. В основе сжатия данных используется свойство концентрации энергии речевых сигналов в малой доле частотной полосы, что позволяет использовать аппроксимацию вида



$$\hat{\vec{x}} = a \sum_{r \in R_m} \vec{x}_r, \quad (3)$$

где

$$\vec{x}_r = A_r \cdot \vec{x}, \quad (4)$$

а A_r – субполосная матрица, определяемая элементами:

$$A_r = \{a_{i,k}^r\}, \quad i, k = 1, \dots, N; \quad a_{ik}^r = \frac{\sin(V_{2,r}(i-k)) - \sin(V_{1,r}(i-k))}{\pi(i-k)}; \quad i \neq k.$$

Коэффициент a при сумме (3) выбирается из условия $\|\hat{\vec{x}}_r\|^2 = m\|\vec{x}\|^2$, что дает:

$$a = \frac{\sqrt{m}\|\vec{x}\|}{\left\|\sum_{r \in R_m} \vec{x}_r\right\|}, \quad (5)$$

где R_m – множество частотных интервалов минимальной суммарной ширины, для которых выполняются условия

$$\sum_{r \in R_m} P_r(\vec{x}) \cong m \|\vec{x}\|^2; \quad (6)$$

$$0,85 \leq m \leq 0,98; \quad (7)$$

где m – множество информационных частотных интервалов.

Соответствующие некоторым частотным интервалам компоненты вида (4) обладают свойством оптимальности в смысле:

$$\int_{\omega \in V_r} |X(\omega) - X_r(\omega)|^2 d\omega + \int_{\omega \notin V_r} |X_r(\omega)|^2 d\omega / 2\pi = \min$$

т.е. наилучшей аппроксимацией отрезков трансформанты Фурье $X_r(\omega)$ исходного вектора в соответствующих частотных интервалах и допускают представление вида

$$\vec{x}_r = \sum_{i=1}^{J_r} \lambda_i^r \alpha_{ir} \vec{q}_i^r; \quad \alpha_{ir} = (\vec{x}, \vec{q}_i^r), \quad (8)$$

где λ_i^r – собственные числа собственных векторов \vec{q}_i^r субполосной матрицы, принимающие значения $0 < \lambda_i^r \leq 1$. Подстановка (8) в (3) дает разложение по набору собственных векторов:

$$\hat{\vec{x}} = \sum_{r \in R_m} \sum_{i=1}^J \beta_{ir} \vec{q}_i^r, \quad (9)$$

где

$$\beta_{ir} = a \lambda_i^r \alpha_{ir}, \quad i = 1, \dots, J_r. \quad (10)$$

Так как наборы собственных векторов \vec{q}_i^r предполагаются известными, то для восстановления исходного отрезка достаточно сохранять информацию о соответствующих коэффициентах разложения. Проведенные исследования показали, что мощность множества частотных интервалов ($\text{int } R_m$) почти для всех звуков русской речи удовлетворяет соотношению

$$\text{int } R_m \approx 0,3R. \quad (11)$$

Поэтому с учетом равенства $J \cong N/R$ получаем коэффициент сжатия за счет использования аппроксимации (3) (по количеству сохраняемых чисел):

$$CH = N / (\text{int } R_m \cdot J) \approx 3. \quad (12)$$



Следующий шаг заключается в применении к коэффициентам разложения квантования по уровню с малым их количеством. В общем виде процедура квантования описывается следующим образом, если выполняется условие:

$$\beta_{ir} \in \hat{O}_m = [\varphi_{m-1}, \varphi_m), \quad (13)$$

то положим

$$\beta_{ir}^* = d_m, m = 1, \dots, K, \quad (14)$$

где K – количество используемых уровней квантования.

Проблема заключается в оптимальном выборе границ отрезков в (13) и значений d_m в (14) в смысле минимизации погрешностей аппроксимации исходных данных квантованными значениями:

$$\varepsilon_i^2 = \sum_{m=1}^K \sum_{\beta_{ir} \in S_m} (\beta_{ir} - d_m)^2, (i = 1, \dots, J), r \in R_m, \quad (15)$$

где S_m – множество значений β_{ir} , удовлетворяющих условию (13). В результате проведенных исследований было показано, что при заданных отрезках \hat{O}_m в (13) минимум правой части (15) достигается на множестве уровней квантования, равных соответствующим средним значениям:

$$d_m = \sum_{\beta_{ir} \in S_m} (\beta_{ir} / \text{int } S_m), m = 1, \dots, K, \quad (16)$$

где $\text{int } S_m$ – мощность множества S_m (число попадающих в них значений β_{ir}).

Введем положительную неубывающую последовательность:

$$0 \leq z_k < z_{k+1}, k = 1; 2 \dots NK - 1, \quad (17)$$

$$NK = J \bullet \text{int } R_m, \quad (18)$$

причем

$$z_k \in \{|\beta_{ir}|\} / \gamma, i = 1, \dots, J; r \in R_m;$$

$$\gamma = \max |\beta_{ir}|, \forall i \text{ и } r \in R_m. \quad (19)$$

Показано, что выполнение условий

$$\sum_{m=1}^K I_m \bar{z}_m^2 = \max, \quad (20)$$

$$\sum_{m=1}^K I_m = NK, \quad (21)$$

где

$$\bar{z}_m = \sum_{i=1}^{I_m} z_{L_{m-1}+i} / I_m, \quad (22)$$

$$L_m = \sum_{i=1}^{m-1} I_i, L_0 = 0,$$

а также выбор уровней квантования в виде

$$\hat{d}_m = \bar{z}_m, m = 1, \dots, K, \quad (23)$$

дает минимум погрешности аппроксимации z_k квантованными значениями

$$z_k^* = \hat{d}_m, \quad (24)$$

когда выполняется условие



$$z_{L_{m-1}+1} \leq z_k \leq z_{L_{m-1}+I_m}. \quad (25)$$

Реально вместо операции (24) следует использовать кодирование

$$\text{cod}z_k = \log_2 m, \quad (26)$$

имея в виду, что номера уровней квантования целесообразно обозначать двоичными числами разрядности p , так что

$$K = 2^p. \quad (27)$$

Таким образом, количество уровней квантования целесообразно выбирать из множества (2; 4; 8 ...). В соответствии с этим в процессе проведения исследований был разработан алгоритм решения задачи (20), (21) с последовательным делением подпоследовательностей на две части, каждая из которых удовлетворяет этим условиям со своими параметрами I_m и d_m (так как деление на две последовательности любой длины несложно реализуется последовательным перебором). Использование стандартизованных последовательностей вида (17)-(19) позволяет не хранить значения уровней (23), а для восстановления данных использовать уровни из заранее сформированной кодовой книги, удовлетворяющей условию

$$\sum_{m=1}^K I_m \cdot (\bar{z}_m - d_m^l)^2 = \min \forall D_l, \quad (28)$$

где

$$D_l = \{d_1^l, \dots, d_K^l\}, d_1^l < d_2^l < \dots < d_K^l. \quad (29)$$

Такие кодовые книги сформированы при $K = 2; 4; 8$ с учетом всех звуков русской речи с усреднением по множеству дикторов. В таблицах 1-3 представлены некоторые из полученных кодовых книг, а также среднее значение уровней квантования для всех звуков и среднеквадратическое отклонение (СКО), вычисляемое по формуле

$$\sigma = \sqrt{\frac{1}{L} \sum_{i=1}^L (Nk_i - \overline{Nk})^2}, \quad (30)$$

где L – количество звуков русской речи; Nk уровень квантования; \overline{Nk} – среднее значение уровня квантования.

Таблица 1

Table 1

Кодовые книги при квазиоптимальном квантовании на 2 уровня
Code books for quasi-optimal quantization of 2 levels

Звуки речи	1-ый уровень	2-ой уровень
А	0.2875	0.7065
Б	0.2138	0.7845
В	0.2262	0.7623
Г	0.2437	0.7497
Д	0.2277	0.7749
Е	0.2425	0.7465
Ё	0.2507	0.7676
Ж	0.2852	0.7143
З	0.2347	0.7726
И	0.2099	0.7838
Й	0.2212	0.7451
К	0.3247	0.6721
Л	0.224	0.7656
М	0.2104	0.7801
Н	0.1955	0.794



Окончание табл. 1

О	0.2801	0.7313
П	0.2904	0.7223
Р	0.2839	0.7168
С	0.3282	0.6911
Т	0.3357	0.6718
У	0.2313	0.7641
Ф	0.3149	0.712
Х	0.2916	0.6782
Ц	0.3155	0.6648
Ч	0.3197	0.68
Ш	0.336	0.6802
Щ	0.3209	0.6771
Ы	0.2335	0.7397
Э	0.2672	0.7201
Ю	0.2354	0.7534
Я	0.2681	0.7262
Среднее значение	0,2661	0,7306
СКО	0,0434	0,0396

Таблица 2

Table 2

Кодовые книги при квазиоптимальном квантовании на 4 уровня

Code books for quasi-optimal quantization of 4 levels

Звуки речи	1-ый уровень	2-ой уровень	3-й уровень	4-ый уровень
А	0.0911	0.3554	0.9133	0.9133
Б	0.0444	0.2847	0.9597	0.9597
В	0.0486	0.3058	0.952	0.952
Г	0.0558	0.3213	0.9344	0.9344
Д	0.0533	0.3003	0.9528	0.9528
Е	0.0558	0.3256	0.9481	0.9481
Ё	0.0654	0.342	0.9453	0.9453
Ж	0.09	0.355	0.9097	0.9097
З	0.0544	0.306	0.95	0.95
И	0.0356	0.2847	0.965	0.965
Й	0.0522	0.2983	0.946	0.946
К	0.1266	0.3917	0.8778	0.8778
Л	0.0431	0.2972	0.9543	0.9543
М	0.0381	0.2779	0.9619	0.9619
Н	0.0303	0.2602	0.9641	0.9641
О	0.0788	0.3603	0.929	0.929
П	0.1028	0.3744	0.8969	0.8969
Р	0.0832	0.3495	0.9148	0.9148
С	0.1239	0.3907	0.8799	0.8799
Т	0.1471	0.3991	0.8557	0.8557
У	0.0563	0.3112	0.9523	0.9523
Ф	0.1191	0.3758	0.8925	0.8925
Х	0.1012	0.3631	0.8983	0.8983
Ц	0.1228	0.3841	0.8587	0.8587
Ч	0.1321	0.3921	0.8705	0.8705
Ш	0.1327	0.3991	0.8708	0.8708
Щ	0.1235	0.3857	0.8641	0.8641
Ы	0.0495	0.3093	0.9488	0.9488



Окончание табл. 2

Э	0.0666	0.342	0.9272	0.9272
Ю	0.0599	0.3399	0.9399	0.9399
Я	0.075	0.3469	0.9116	0.9116
Среднее значение	0,0793	0,3396	0,9208	0,9208
СКО	0,0346	0,04047	0,0353	0,0353

Таблица 3

Table 3

Кодовые книги при квазиоптимальном квантовании на 8 уровня

Code books for quasi-optimal quantization of 8 levels

Звуки речи	1-ый уровень	2-ой уровень	3-й уровень	4-ый уровень	51-ый уровень	6-ой уровень	7-ой уровень	8-ой уровень
А	0.0134	0.1334	0.3973	0.3973	0.6254	0.6697	0.8711	0.9878
Б	0.005	0.0722	0.3207	0.3207	0.6998	0.7408	0.9327	0.9962
В	0.008	0.0788	0.3452	0.3452	0.6747	0.7172	0.9195	0.9915
Г	0.0061	0.0881	0.3621	0.3621	0.6543	0.6962	0.902	0.9932
Д	0.0079	0.083	0.3401	0.3401	0.6926	0.7324	0.9256	0.9943
Е	0.0056	0.0887	0.3671	0.3671	0.6419	0.6849	0.9148	0.994
Ё	0.0043	0.103	0.3877	0.3877	0.6657	0.6993	0.9139	0.9962
Ж	0.0147	0.1304	0.396	0.396	0.6291	0.6681	0.8705	0.987
З	0.0091	0.0832	0.3462	0.3462	0.6823	0.719	0.924	0.9937
И	0.0032	0.0599	0.3204	0.3204	0.6954	0.7397	0.9406	0.9974
Й	0.0065	0.082	0.3371	0.3371	0.6579	0.7112	0.9096	0.9937
К	0.0249	0.1745	0.4354	0.4354	0.5901	0.6226	0.8304	0.9785
Л	0.0044	0.0712	0.3362	0.3362	0.6793	0.7236	0.9238	0.9954
М	0.0049	0.0642	0.3125	0.3125	0.6967	0.742	0.9371	0.9969
Н	0.0019	0.0525	0.2931	0.2931	0.7156	0.7623	0.9411	0.9979
О	0.0111	0.1217	0.404	0.404	0.6395	0.6772	0.8875	0.9912
П	0.0173	0.1494	0.4202	0.4202	0.6111	0.6457	0.8485	0.9826
Р	0.0123	0.12	0.3906	0.3906	0.6333	0.6735	0.878	0.9874
С	0.0279	0.1672	0.4297	0.4297	0.6157	0.6443	0.8376	0.973
Т	0.0367	0.1953	0.4419	0.4419	0.5881	0.6148	0.8059	0.9632
У	0.0054	0.0907	0.3538	0.3538	0.6679	0.7113	0.9209	0.9957
Ф	0.0238	0.166	0.4194	0.4194	0.6291	0.6589	0.8497	0.98
Х	0.0189	0.146	0.4019	0.4019	0.5885	0.628	0.8526	0.9809
Ц	0.0278	0.1644	0.4264	0.4264	0.5786	0.6111	0.8101	0.9666
Ч	0.0298	0.1846	0.4346	0.4346	0.5867	0.6206	0.82	0.9702
Ш	0.0304	0.1859	0.4425	0.4425	0.5984	0.6287	0.8213	0.9731
Щ	0.0241	0.1695	0.4315	0.4315	0.5926	0.627	0.8154	0.9724
Ы	0.0049	0.0816	0.3489	0.3489	0.6466	0.6964	0.9171	0.9964
Э	0.009	0.1052	0.3895	0.3895	0.6251	0.6709	0.8852	0.9908
Ю	0.0065	0.0971	0.3805	0.3805	0.6402	0.6866	0.8939	0.9925
Я	0.0112	0.111	0.394	0.394	0.6288	0.6717	0.8666	0.9897
Среднее значение	0,0134	0,1168	0,3808	0,3808	0,641	0,6805	0,8828	0,9871
СКО	0,0097	0,0427	0,0429	0,0429	0,0389	0,0431	0,0434	0,0102



Для иллюстрации работоспособности и эффективности разработанных метода и алгоритма были проведены экспериментальные исследования, которые показали, что разборчивость речи сохраняется уже при $K = 2$. Таким образом, с учетом необходимости сохранения знакового разряда и значения γ достигаемый максимальный коэффициент сжатия может быть равен

$$CH_{\max} = 12N / (N + 12), \quad (31)$$

(в предположении 8-разрядности исходных отсчетов). То есть если N достаточно велико, то

$$CH_{\max} \cong 12. \quad (32)$$

Заключение

В результате проведенных исследований было установлено, что предлагаемый метод сжатия речевых данных без пауз на основе оптимального квантования по уровню коэффициентов разложения отрезков речевых сигналов по собственным векторам субполосных матриц из m -информационных частотных интервалов с применением кодовых книг квазиоптимальных квантователей позволяет в зависимости от величины разрядности исходных отсчетов указанных сигналов обеспечить коэффициент сжатия до 12 раз.

Выводы

Использование разработанного метода позволит достичь общего коэффициента сжатия речевых данных как за счет обнаружения и кодирования пауз, которые могут составлять более 60% от продолжительности диалога, так и за счет квантования по уровню коэффициентов разложения отрезков речевых сигналов по собственным векторам субполосных матриц из m -информационных частотных интервалов величин в 20-25 раз.

Исследования выполнены при поддержке гранта РФФИ № 15-07-01463.

Список литературы

References

1. Ashwin R., Kumaresan R. 2000. On decomposing speech into modulated components. IEEE Transactions on Speech and Audio Processing, 8(3): 240–254.
2. Seki H. 1958. A new method of speech transmission by frequency division and multiplication. The Journal of the Acoustical Society of Japan, 14: 138–142.
3. Алдошина И. 2002. Основы психоакустики. Слух и речь. Часть 1. Информационно-технический журнал «Звукорежиссер». 1: 38–44.
4. Aldoshina I. 2002. Fundamentals of psychoacoustics. Rumor and speech. Part 1. Information and Technical Journal "Sound producer". 1: 38–44.
5. Белов С.П., Прохоренко Е.И. 2006. Об уменьшении объема трафика при пакетной передаче речевых сообщений за счет кодирования пауз. Научные ведомости БелГУ. Сер. Информатика и прикладная математика. 1(21): 141–148.
6. Belov S.P., Prokhorenko E.I. 2006. On the reduction in the volume of traffic for packet transmission of voice messages due to the coding of pauses. Belgorod State University Scientific Bulletin. Informatics and Applied Mathematics. 1(21): 141–148.
7. Быков С.Ф., Журавлев В.И., Шалимов И.А. 2003. Цифровая телефония. М., Радио и связь, 144.
8. Bykov S.F., Zhuravlev V.I., Shalimov I.A. 2003. Digital telephony. M., Radio and Communication, 144.
9. Гусинская Е.И., Зайцев А.А. 2004. Оптимизация банка фильтров в задачах субполосного кодирования: тематический обзор. Научно-технический журнал Цифровая обработка сигналов. 3(12): 18–29.
10. Gusinskaya E.I., Zaitsev A.A. 2004. Optimization of the filter bank in sub-band coding problems: a thematic review. Scientific and technical journal Digital signal processing. 3(12): 18–29.



7. Жилияков Е.Г., Белов С.П., Прохоренко Е.И. 2005. О сжатии речевых сигналов. Вестник Национального технического университета «ХПИ». Сборник научных трудов. Тематический выпуск: Информатика и моделирование. Харьков: 32-41.

Zhilyakov E.G., Belov S.P., Prokhorenko E.I. 2005. On the compression of speech signals. Bulletin of the National Technical University "KhPI". Collection of scientific papers. Thematic issue: Informatics and modeling. Kharkov, 56: 32-41.

8. Жилияков Е.Г., Белов С.П., Прохоренко Е.И. 2007. Методы обработки речевых данных в информационно-телекоммуникационных системах на основе частотных представлений. Белгород, Изд-во БелГУ, 138.

Zhilyakov E.G., Belov S.P., Prokhorenko E.I. 2007. Methods of processing speech data in information and telecommunication systems on the basis of frequency representations. Belgorod, Publishing house of BelSU, 138.

9. Жилияков Е.Г., Белов С.П., Прохоренко Е.И. 2007. Новый метод вычисления спектрограмм речевых сигналов. Сборник докладов 13-й Междунар. науч.-техн. конф. «Радиолокация, навигация, связь». Воронеж, 1: 278-283.

Zhilyakov E.G., Belov S.P., Prokhorenko E.I. 2007. A new method for calculating the spectrograms of speech signals. Collected papers of the 13th Intern. Scientific-techn. Conf. "Radar, navigation, communication". Voronezh, 1: 278-283.

10. Жилияков Е.Г., Белов С.П., Прохоренко Е.И. 2007. Уменьшение объема битового представления речевых данных на основе нового метода удаления пауз. Вопросы радиоэлектроники. Серия Электронная вычислительная техника (ЭВТ). М., 2: 124-139.

Zhilyakov E.G., Belov S.P., Prokhorenko E.I. 2007. Decrease in bit representation of speech data on the basis of a new method of deleting pauses. Questions of radio electronics. The series Electronic computing (EWT). М., 2: 124-139.

11. Жилияков Е.Г., Белов С.П., Белов А.С., Фирсова А.А. 2010. Об эффективности различных подходов к сегментации речевых сигналов на основе обнаружения пауз. Научные ведомости БелГУ. Сер. Информатика. 7(78): 187-193.

Zhilyakov E.G., Belov S.P., Belov A.S., Firsova A.A. 2010. On the effectiveness of different approaches to the segmentation of speech signals based on the detection of pauses. Belgorod State University Scientific Bulletin. Informatics. 7(78): 187-193.

12. Калинцев Ю.К. 1991. Разборчивость речи в цифровых вокодерах. М., Радио и связь, 220.

Kalintsev Yu.K. 1991. Intelligibility of speech in digital vocoders. М., Radio and Communication, 220.

13. Ковалгин Ю.А., Вологдин Э.И. 2004. Цифровое кодирование звуковых сигналов. СПб., Корона-принт, 240.

Kovalgin Yu.A., Vologdin E.I. 2004. Digital coding of sound signals. St. Petersburg, Crown-print, 240.

14. Санников В.Г. 2003. Методы кодирования речевых сигналов: Учебное пособие. М., МТУСИ, 63.

Sannikov V.G. 2003. Methods of encoding speech signals: Textbook. Moscow, MTUCI, 63.

15. Сергиенко А.Б. 2003. Цифровая обработка сигналов. СПб., Питер, 604.

Sergienko A.B. 2003. Digital signal processing. St. Petersburg, Peter, 604.

16. Сергиенко В.С., Баринов В.В. 2009. Сжатие данных, речи, звука и изображений в телекоммуникационных системах. М.: Радио Софт, 360.

Sergienko V.S., Barinov V.V. 2009. Compression of data, speech, sound and images in telecommunication systems. М.: Radio Soft, 360.

17. Сжатие данных в системах сбора и передачи информации. Под ред. В.А. Свириденко. 1985. М.: Радио и связь, 184.

Compression of data in information collection and transmission systems. Ed. V.A. Sviridenko. 1985. Moscow: Radio and Communication, 184.

18. Синильников А.М. 1988. Полосовое кодирование звуковых сигналов с ортогональным преобразованием. Электросвязь. 9: 34-36.

Sinilnikov A.M. 1988. Band coding of audio signals with orthogonal transformation. Telecommunications. 9: 34-36.

19. Сэломон Д. 2004. Сжатие данных, изображений и звука. М., ТЕХНОСФЕРА, 368.

Salomon D. 2004. Compression of data, images and sound. М., TECHNOSPHERE, 368.

20. Цифровая обработка и передача речи. Под ред. О.И. Шелухина. 2000. М., Радио и связь, 456.

Digital processing and voice transmission. Ed. O.I. Shelukhina. 2000. Moscow, Radio and Communication, 456.